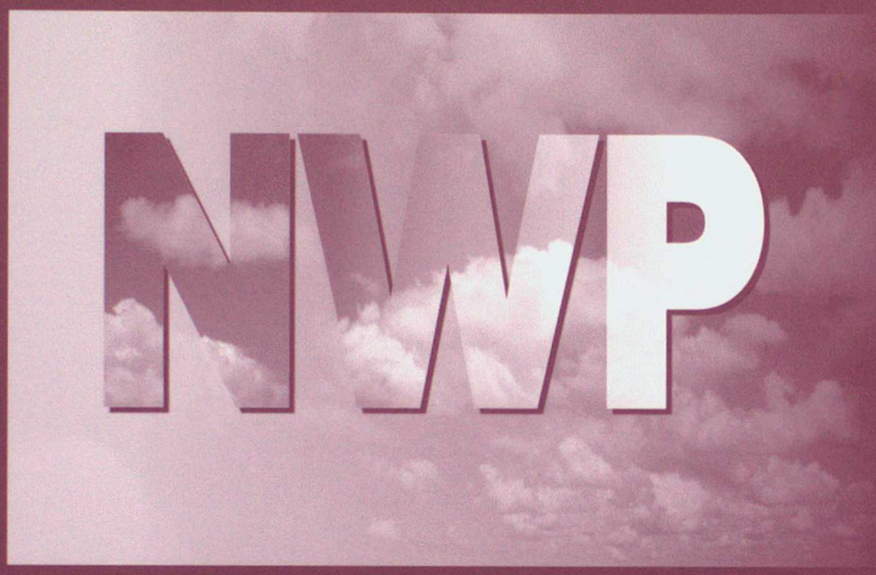
Numerical Weather Prediction



Forecasting Research Scientific Paper No. 71

A 3D variational data assimilation system for the stratosphere and troposphere

By R. Swinbank, N.B. Ingleby, P.M. Boorman and R.J. Renshaw

July 2002

© Crown Copyright 2002

Met Office , NWP Division , Room 344, London Road , Bracknell , Berkshire ,RG12 2SZ, United Kingdom

Numerical Weather Prediction

Forecasting Research Scientific Paper No. 71

A 3D variational data assimilation system for the stratosphere and troposphere

By R. Swinbank, N.B. Ingleby, P.M. Boorman and R.J. Renshaw

NWP Division
Met Office
London Road
Bracknell
Berkshire
RG12 2SZ
United Kingdom

© Crown Copyright 2002

Permission to quote from this paper should be obtained from the Director NWP, at the above address.

For additional copies, or information about other NWP publications, contact:
john.armstrong@metoffice.com; tel: +44 (0)1344 856245; fax: +44 (0)1344 854026

Approved for issue as a Scientific Paper by A.C. Lorenc, Head (DA), 16 July 2002

Last revised by R. Swinbank 19 July 2002

A 3D variational data assimilation system for the stratosphere and troposphere

R. Swinbank, N.B. Ingleby, P.M. Boorman and R.J. Renshaw

Met Office, Bracknell, UK

Abstract

The Met Office has developed a 3D variational (3D-VAR) data assimilation system based on the stratosphere-troposphere configuration of the Unified Model. The stratosphere-troposphere data assimilation system is based on the version used for the standard operational global forecast model, which was implemented in March 1999.

The 3D-VAR assimilation system described here replaced the former Analysis Correction (AC) assimilation system in November 2000. This paper describes both the impacts of moving from the AC to VAR, and of adding additional stratospheric levels to the standard global data assimilation system.

1. Introduction

The aim of a data assimilation system is to construct a self-consistent sequence of three-dimensional analyses of the atmosphere. This sequence of synoptic analyses is constructed from a heterogeneous mixture of observational data, using an atmospheric model to organise and summarise information from those observations. As computing power has increased over the last few decades, the production of meteorological analyses has become ever more sophisticated.

Perhaps the key development in atmospheric data analysis was the adoption of statistical interpolation methods, in which data from observations and from forecast fields are combined by weighting them according to their estimated errors. This approach formed the basis of so-called "optimal interpolation" (OI) data assimilation schemes, which were widely adopted for operational weather forecasting (e.g. *Lorenc* [1981]) once adequate computer power became available. Probably the greatest limitation of those early implementations is that only a limited set of observations were used to calculate the analysed field values at each model grid point. The Analysis Correction scheme [*Lorenc et al,* 1991] was developed at the Met Office to overcome this data selection problem, but it involved other approximations to the solution of the full OI equations.

Since more powerful computers have become available, variational methods have been developed which allow a closer approximation to the statistically optimal solution. The first operational implementation of a variational data assimilation system was the SSI system at NCEP [*Parrish and Derber*, 1992], followed by ECMWF [*Courtier et al*, 1998]. An alternative approach to solving the same basic statistical interpolation equations, known as the Physical-space Statistical Analysis Scheme (PSAS), was developed at the Data Assimilation Office [*Cohn et al*, 1998]. The Met Office three-dimensional variational (3D-VAR) system was first implemented in the standard operational global forecast suite in March 1999 [*Lorenc*

et al, 2000]. Versions of the 3D-VAR system were subsequently developed for other model configurations, most recently for the stratosphere-troposphere model.

Since October 1991 daily stratospheric analyses have been produced at the Met Office using a stratosphere-troposphere data assimilation system [*Swinbank and O'Neill*, 1994]. Initially, the analyses were produced using a special research version of the Analysis Correction (AC) data assimilation scheme [*Lorenc et al*, 1991] to support the Upper Atmosphere Research Satellite project. Subsequently, the system was run as part of the operational forecast suite. On 14th November 2000 the stratospheric data assimilation system was converted to use 3D-VAR.

In the following sections of this paper we describe the stratospheric version of the 3D-VAR system. Section 2 gives an overview of our implementation of 3D-VAR. The basic specification of the background error covariances follows *Ingleby* [2001], but the stratospheric system includes variations on this approach, described in section 3. Before operational implementation of the system, a range of trials was run. Section 4 uses results from these assimilation trials to show the impact of replacing AC by 3D-VAR together with the effect of adding extra stratospheric levels to the standard global forecast models. Section 4 focuses on objective forecast verification statistics, which emphasise objective scores in the troposphere and lower stratosphere, to which virtually all the non-satellite data are restricted. To complement those results, section 5 presents some subjective assessments of the analyses, focused on the middle and upper stratosphere. The paper concludes with a final section, setting out our plans for future data assimilation developments.

2. The variational data assimilation scheme

2.1 Variational Analysis

A variational analysis scheme is formulated in such a way that the optimal analysis is defined by the model state $\mathbf{x}$ which minimises a cost function J. In standard notation [*Ide et al*, 1997], J is made up by a background term, quantifying the difference between the model state and a prior (background) estimate of $\mathbf{x}$, and an observation term, quantifying the difference between observed values and the equivalent background values:

$$J(\mathbf{x}) = \frac{1}{2}\left(\mathbf{x}^b - \mathbf{x}\right)^T \mathbf{B}^{-1}\left(\mathbf{x}^b - \mathbf{x}\right) + \frac{1}{2}\left(\mathbf{y}^o - H(\mathbf{x})\right)^T \left(\mathbf{E}+\mathbf{F}\right)^{-1}\left(\mathbf{y}^o - H(\mathbf{x})\right) \qquad (1)$$

In Equation 1 $\mathbf{x}^b$ is a background estimate of $\mathbf{x}$, with error covariance $\mathbf{B}$; $\mathbf{y}^o$ is a vector of observed values, which have error covariance $\mathbf{E}$; H is an observation operator, which derives an observation value from the model state (this can be considered to be a generalised interpolation), with associated error covariance $\mathbf{F}$.

Following *Courtier et al* [1994], the Met Office 3D-VAR scheme is formulated as an incremental variational problem, in which we solve for a model perturbation $\mathbf{w}'$, rather than the full model state $\mathbf{x}$. In most implementations of the Met Office VAR scheme, the perturbation state $\mathbf{w}'$ is at a lower resolution than the full model, but in the operational stratospheric configuration the resolutions are the same. However, the horizontal and vertical grid staggering of the model variables are changed, as discussed in section 2.2.

For a practical solution of the minimisation problem, it is necessary to improve the conditioning of the Hessian matrix. The Hessian matrix comprises second-order partial derivatives of J with respect to the control variables. Ignoring the second derivatives of H, the Hessian is given by

$$\left(\frac{\partial^2 J}{\partial \mathbf{w}'^2}\right) = \mathbf{B}^{-1} + \mathbf{H}^T (\mathbf{E}+\mathbf{F})^{-1} \mathbf{H} . \tag{2}$$

The linearised generalised interpolation **H** in the second term depends on the observations being used, and it is hard to analyse its conditioning in a general way. Instead, we concentrate on the first term, which depends on the background error covariance. As it stands, $\mathbf{B}^{-1}$ is very ill-conditioned; balanced and smooth modes correspond to small eigenvalues of $\mathbf{B}^{-1}$, while imbalanced or rough modes correspond to large eigenvalues.

To alleviate this ill-conditioning we use a transformation (**U**) to independent modes, together with its generalised inverse $\mathbf{U}^{-1}$. We define these such that approximately:

$$\mathbf{B}^{-1} = \left(\mathbf{U}^{-1}\right)^T \mathbf{U}^{-1} , \tag{3}$$

$$\mathbf{B} = \mathbf{U}\mathbf{U}^T . \tag{4}$$

In other words, **U** is essentially the "square root" of the background covariance **B**, so the background error covariance is defined by the transform **U**.

Defining a new control variable **v**, such that **w'**=**Uv**, the Hessian matrix of the transformed problem can be written:

$$\left(\frac{\partial^2 J}{\partial \mathbf{v}^2}\right) = \mathbf{I} + \mathbf{U}^T \mathbf{H}^T (\mathbf{E}+\mathbf{F})^{-1} \mathbf{H}\mathbf{U} \tag{5}$$

Now that the first term has become the identity matrix, the Hessian has become much better conditioned, so it is possible to find the solution within a reasonable number of iterations.

The transform **U** is itself constructed from three successive transforms: $\mathbf{U}_p$ (parameter transform), $\mathbf{U}_v$ (vertical transform) and $\mathbf{U}_h$ (horizontal transform). The parameter transform $\mathbf{U}_p$ converts between model variables and a set of analysis variables, which are constructed in such a way as to attempt to minimise the correlation between the errors in the different parameters (any remaining cross-correlations are ignored). The first analysis parameter ψ (streamfunction) is intended to represent the balanced flow, and the second parameter χ (velocity potential) the other (irrotational) component of the wind field. The unbalanced pressure A_p is that component of the mass field that is not balanced with the wind field. The humidity variable is relative humidity. The vertical transform $\mathbf{U}_v$ separates each three dimensional field into two-dimensional fields of vertical mode coefficients. The calculation of the vertical modes and the two-dimensional horizontal transform $\mathbf{U}_h$ are described more fully in section 3.

2.2 Stratosphere-troposphere implementation

The new stratosphere-troposphere data assimilation system is based on a 40 level (40-L) version of the Unified Model [*Cullen et al*, 1993] with its top around 0.1 hPa, and a vertical resolution of about 2.5 km in the stratosphere, see Figure 1. The model levels were chosen so that they are the same as the standard 30-level (30-L) global forecast model in the troposphere, but extra stratospheric levels were added, with the aim of allowing improved processing of satellite temperature soundings. The vertical resolution in the stratosphere is coarser than in the 42-level (42-L) model used by the old stratosphere-troposphere assimilation system. This decision was made because a major aim of the new system was to act as a prototype for a possible future extended global data assimilation system for operational forecasting; in that context it was thought difficult to justify high vertical resolution. The horizontal resolution is 2.5° latitude by 3.75° longitude (exactly the same as for the old stratosphere-troposphere system and the climate model; also referred to as N48 from the number of grid lengths in 180° longitude). Flow features in the stratosphere tend to be of larger scale than in the troposphere, essentially because smaller scale Rossby waves cannot propagate up through the stratosphere [*Charney and Drazin*, 1961], so this coarse resolution is sufficient to capture the main features of the stratospheric circulation.

The model uses a set of physical parametrizations similar to those used by the HadAM3 version of the climate model, and similar to that described by *Swinbank et al* [1999] and *Jackson et al* [2001]. Of particular importance in the stratosphere is the specification of the orographic gravity wave drag scheme [*Milton and Wilson*, 1996]. This includes the effects of anisotropic orography and enhanced low-level wave breaking compared to the original *Palmer et al* [1986] scheme. The gravity-wave drag scheme is replaced by Rayleigh Friction above 20 hPa. Although this model does not include a parametrization of non-orographic gravity wave drag, the Rayleigh Friction is intended to be a crude representation of the effect of both orographic and non-orographic gravity waves in the upper stratosphere and lower mesosphere. The radiation scheme is due to *Edwards and Slingo* [1996], which includes the effect of a range of important trace gases. Comparisons with line-by-line calculations show that the Edwards-Slingo radiation scheme is performing well in both short and long wave regimes throughout the stratosphere, with some discrepancies in the mesosphere. Other physical parametrizations include the MOSES surface exchange scheme [*Cox et al*, 1999]

The current version of the Unified Model uses a pressure-based hybrid vertical coordinate, with an unstaggered (Lorenz) vertical grid and an Arakawa 'B' horizontal grid. However, the Met Office VAR scheme has been designed to use the grid and variables of the "New Dynamics" scheme being developed for future versions of the Unified Model [*Malcolm et al*, 2001]. The new dynamical core is non-hydrostatic and uses a height-based vertical coordinate, with Charney-Phillips vertical staggering and an Arakawa 'C' horizontal grid. Thus the transformations between the model state $\mathbf{x}$ and the perturbation state $\mathbf{w}'$ include interpolation between the two different grid systems. Partly as a result of this, the uninitialised analysis is insufficiently balanced, and unsuitable for starting a forecast.

For simplicity, and because of its similarity with the repeated insertions used in the previous (AC) scheme, we use an Incremental Analysis Update (IAU) scheme [*Bloom et al*, 1996] to initialise the model. The calculated analysis increments are

added gradually to the model over a 6-hour period centred on the nominal analysis time.

2.3 Assimilation of satellite data

Another major development has been in the assimilation of satellite soundings. The adoption of 3D-VAR has made it possible to assimilate radiances from TOVS and ATOVS directly, while the previous (AC) system used temperature retrievals derived by NESDIS from TOVS measurements. In the standard global assimilation system the introduction of ATOVS data from NOAA-15 [*English et al*, 2000] was simultaneous with the adoption of 3D-VAR and gave a broadly similar benefit. Subsequently, the direct assimilation of radiances (implemented October 1999) also gave significant benefit. The use of NOAA-16 ATOVS data was also beneficial; those data were introduced into both standard global and stratospheric assimilation system in April 2001, after the implementation of 3D-VAR for the stratosphere.

One problem with the standard 30-level global model is that the model domain does not span the range of levels measured by the ATOVS sounder; the uppermost channels are above the model top. Thus the radiance measurements from the upper channels can only be used by extrapolating the temperature structure above the model top (the alternative would be to ignore all channels with significant contributions above the top model level). In deciding the vertical configuration of the stratosphere-troposphere model, an important consideration was to ensure that the model top is sufficiently high to cover the weighting functions of all the ATOVS sounder channels. The improved stratospheric resolution has also made it possible to use data from some higher channels that are not used in the 30-level global model, over high land.

In order to be able to get the most beneficial impact from satellite data, it is important to take account of small but significant biases between satellite radiance measurements and the equivalent radiances calculated by the forward model. These biases are often as large as the signal in the radiances. We correct these biases in two parts (e.g. *Harris and Kelly* [2001]). Statistics of observed minus model background radiance are accumulated over a week or more, for each channel and for each scan position (viewing angle). The global mean differences form the first part of the bias correction. There remains a residual component of bias that varies on synoptic scales. We model this as a linear combination of two "airmass" predictors. The predictors we choose are the observed radiances in AMSU channels 5 and 9. Linear regression on the accumulated statistics provides coefficients for these predictors. The linear combination of coefficients times predictors forms the second part of the bias correction.

Although the bias corrections can be applied successfully in the troposphere and lower stratosphere, the same approach cannot be used in the middle and upper stratosphere. The reason is that the biases in the model background fields are relatively small in the troposphere, and what model biases there are kept to a minimal level by assimilation of well-characterised observations such as radiosonde soundings. Above the lower stratosphere, satellite soundings are essentially the only temperature observations, and the biases in the model are themselves relatively large. Applying bias corrections to the satellite data would tend to make the satellite data track the model biases. We have thus decided not to bias-correct those satellite channels that

peak above the lower stratosphere (AMSU channels 10-14 on ATOVS) and only partially correct AMSU channel 9, which peaks in the lower stratosphere.

3. Background error covariances

3.1 Overview

As with the standard 30-level (or 30-L) global forecast model, the background (forecast) error covariances were calculated using the so-called "NMC method" [*Parrish and Derber*, 1992], using differences between T+48 and T+24 forecast fields valid at the same time. These forecasts were taken from the 40-L Analysis Correction control runs and included 60 cases in total. The resulting error covariances are then used to derive the vertical ($\mathbf{U}_v$) and horizontal ($\mathbf{U}_h$) transforms. The background error covariances are represented using vertical modes (derived from a globally averaged vertical covariance matrix) each with its own horizontal correlation function. This approach allows (partially) for inhomogeneities in the horizontal: vertical correlations are narrower in the tropics, but there is also a variation linked with tropopause height in the extratropics. In the standard global 3D-VAR the association of short (long) vertical scales with short (long) horizontal scales was modelled reasonably well, but the longer horizontal scales in the stratosphere were largely absent. This has been addressed using localised (rotated) vertical modes in the stratosphere. *Ingleby* [2001] describes the background error covariances used for the original Met Office 3D-VAR system. But since the rotation of vertical modes and a modification to the horizontal spectra (both implemented in the operational 30-L global system on 16 May 2000) were not described in that paper, further details are given here.

3.2 Vertical modes and their rotation

Vertical covariances are calculated in 5° latitude bands; these are combined over different cases and summed to form an area-weighted global covariance matrix $\mathbf{B}_v$ for each variable. We perform an eigendecomposition of $\mathbf{B}_v$, weighted by the diagonal matrix $\mathbf{P}_v$ – which gives an inner product proportional to the global mean mass in each layer. (Another possibility is to set the diagonal elements of $\mathbf{P}_v$ to the reciprocals of the standard deviations, thus performing a decomposition of the correlation matrix):

$$\mathbf{P}_v\mathbf{B}_v\mathbf{P}_v = \mathbf{F}\Lambda\mathbf{F}^T. \tag{6}$$

Λ is a diagonal matrix of eigenvalues and $\mathbf{F}\mathbf{F}^T = \mathbf{I}$. However any factorisation $\mathbf{P}_v\mathbf{B}_v\mathbf{P}_v = \mathbf{A}\mathbf{M}\mathbf{A}^T$ could be used to represent $\mathbf{B}_v$. In particular we can use

$$\mathbf{P}_v\mathbf{B}_v\mathbf{P}_v = \mathbf{F}\Lambda^{\frac{1}{2}}\mathbf{R}\mathbf{R}^T\Lambda^{\frac{1}{2}}\mathbf{F}^T, \tag{7}$$

where $\mathbf{R}$ is a rotation matrix i.e. $\mathbf{R}\mathbf{R}^T = \mathbf{I}$. We choose $\mathbf{R}$ to give a 'varimax' rotation (see, e.g., *Richman* [1986]), which tends to give a localised (perhaps more physical) representation. To model the latitudinal and seasonal variation of vertical covariances we project the local (5°) covariances, $\mathbf{B}_i$, onto the global eigenvectors $\mathbf{F}_v$. That is we calculate $\mathbf{M}_i$ from

$$\mathbf{B}_i = \mathbf{W}\mathbf{M}_i\mathbf{W}^T, \tag{8}$$

where

$$\mathbf{W} = \mathbf{P}_v^{-1}\mathbf{F}\boldsymbol{\Lambda}^{\frac{1}{2}}\mathbf{R}. \tag{9}$$

If $\mathbf{B}_i = \mathbf{B}_v$ then $\mathbf{M}_i = \mathbf{I}$. In general $\mathbf{M}_i$ will not be diagonal but we assume that it is approximately so and discard the off-diagonal elements. If desired the representation can be truncated before rotation (i.e. discard the eigenvectors corresponding to the smallest eigenvalues), but we do not currently do this.

Note that as far as the 'global' decomposition is concerned $\mathbf{B}_v$ is represented exactly – with or without rotation. However, for a particular latitude band, use of rotated modes will provide a different approximation to the original $\mathbf{B}_i$. We also have an option to replace streamfunction (velocity potential) vertical modes with those for rotational (divergent) kinetic energy. The KE modes are narrower and this reduces the implied vertical scales, they also slightly improve the positioning of variance maxima associated with mid-level jets. This option was implemented in our operational global 3D-VAR in July 1999 (some results are given in *Ingleby* [2001]).

There is a lot of freedom in how $\mathbf{B}_v$ is decomposed. Initial trials suggested that rotation is beneficial in the stratosphere but not in the troposphere. (In the troposphere use of rotated modes gives a reduction of scale selectivity - implied vertical correlations change less as a function of latitude and horizontal scale.) The matrix $\mathbf{W}$ was decomposed into tropospheric and stratospheric modes $\mathbf{W}_A$ and $\mathbf{W}_B$ respectively, so

$$\mathbf{W} = \left(\mathbf{W}_A \mathbf{W}_B\right), \tag{10}$$

$$\mathbf{B}_v = \mathbf{W}\mathbf{W}^T = \mathbf{W}_A\mathbf{W}_A^T + \mathbf{W}_B\mathbf{W}_B^T. \tag{11}$$

We then reorthogonalise $\mathbf{P}_v\mathbf{W}_A\mathbf{W}_A{}^T\mathbf{P}_v$, as in equation 6.

Systems representing vertical covariance as a function of horizontal wavenumber (e.g. *Derber and Bouttier*, [1999]) can automatically include longer stratospheric scales, but they have less variation with latitude. The NAVDAS system [*Daley and Barker*, 2000] also had a problem representing longer stratospheric scales. Their solution was to modify $\mathbf{P}_v$ in eqn. 1 so that the first mode was largely stratospheric. None of the methods currently in use model the full 3-Dimensional structure.

3.3 Horizontal Spectra

We assume that the horizontal correlations (of the coefficients of vertical modes) are homogeneous and isotropic – it is then convenient to convert to a spectral representation because the spectral coefficients are statistically independent [*Boer* 1983; *Courtier et al* 1998]. The spectrum of stream function horizontal covariance C is given by $D_n(\psi)$ where n is the global wavenumber. The differential length scale $\lambda(\psi)$ is

$$\lambda^2(\psi) = \frac{-2C(0)}{\nabla^2 C(0)} = \frac{2a^2\sum D_n(\psi)}{\sum D_n(\psi)n(n+1)}, \tag{12}$$

where a is the radius of the Earth and summations are over n. Similarly (although it doesn't have a corresponding covariance function) we define the Rotational Kinetic Energy length scale

$$\lambda^2(RKE) = \frac{2a^2 \sum D_n(\psi) n(n+1)}{\sum D_n(\psi) n^2 (n+1)^2} . \tag{13}$$

The original 3D-VAR implementation used a Second Order Auto-Regressive (SOAR) function.

$$f(r) = (1 + r/d) e^{-r/d}, \tag{14}$$

where r is distance and for unbalanced pressure and relative humidity (RH) the distance parameter d is taken as $s\lambda$ using the appropriate λ and a tuning factor s ($=1/\sqrt{2}$). For stream function we used $d(\psi) = s\lambda(RKE)\sqrt{3}$ and similarly for velocity potential. (The factor of $\sqrt{3}$ was intended to match up original and used RKE (DKE) length scales, it arises from the fourth derivative of the SOAR in one-dimension. However in two-dimensions the fourth derivative of a SOAR becomes infinite at the origin, and hence the RKE length scale goes to zero as resolution increases. Alternatives to this usage are being considered.) We then calculate the spectra D_n^{SOAR} from $f(r)$ numerically. Note that the scales λ calculated from D_n^{SOAR} are slightly longer than the input scales d, particularly for short d, this is primarily due to the fact that the Legendre transform used has poor resolution near the poles. The tuning factor s and the use of KE length scales was dominant however and in general the implied length scales were shorter than the original ones, considerably so for ψ and χ. The shortening was deliberate as the original scales seemed rather large, and for ψ and χ the significant variance at small n has little effect on wind correlations.

As discussed in section 1 of *Ingleby* [2001] the NMC-method may give length scales that are too long, however use of the SOAR functions as described above may have reduced them too much. One option tried was modifying the correlations from the original spectra to make them 'compactly supported' (see *Gaspari and Cohn* [1999] and p 1790 of *Courtier et al* [1998]). However the solution eventually adopted was to blend the original and SOAR spectra.

$$D_n^{NEW} = \left\{ D_n D_n^{SOAR} \right\}^{0.5} \tag{15}$$

As expected this gives intermediate results, and the correlations become very small at large distance although not formally compactly supported. The resulting correlation function is the convolution of the two functions which when convolved with themselves give the original correlation and the SOAR respectively (see p 1787 and 1790 of *Courtier et al* [1998]). For further details see the Annex to *Ingleby* [2002].

3.4 Covariances used and implied

Figure 2 shows vertical modes (columns of **W**) for rotational kinetic energy (used for ψ); the rotational wind control variable determines most of the wind and (via balance relationships) mass covariances in the extratropics. As described in section 3.2 the modes were rotated and the 19 "highest" rotated modes are retained, the

remainder being reorthogonalised (this was also applied to DKE/χ and A_p modes but not RH). The top left panel shows the first six reorthogonalised modes – these are similar to the first six modes before rotation except that amplitudes above about level 27 are very small. The first mode, which has the same sign at all levels and a maximum near the tropopause, represents 'equivalent barotropic' flow (*Ingleby* [2001], p 221); the others show increasing numbers of sign changes (typical of eigenvectors). The top right and bottom left panels show some of the rotated modes, these are clearly more localised, largely of the same (arbitrary) sign, and there is a tendency for them to have similar shapes and scales. The bottom right panel shows the weighted variance explained by each of the modes.

Horizontal length scales calculated from the forecast differences are given in table 1 for selected levels. In general they are longer than the corresponding scales for the 30-L version shown in table 1 of *Ingleby* [2001] – this is most likely due to the longer scales used in the AC scheme used to generate the 40-level statistics. (From experience modifying the horizontal spectra, longer scales tend to give worse verification against analyses at short range, but the performance at longer range or against observations cannot be inferred.) However at about 30 hPa most of the 40-L length scales are shorter than the corresponding 30-L scales and A_p and T scales decrease above this level. The coarse vertical resolution in the 30-L stratosphere is probably responsible for excessively long horizontal scales; these scales are reduced by the extra 40-L stratospheric levels and make the contrast between tropospheric and stratospheric scales less dramatic, although still substantial.

Table 2 shows implied length scales at N48 horizontal resolution. The stratospheric length scales are longer than the corresponding tropospheric ones (not the case for the original unrotated vertical modes – see Table 2 of *Ingleby* [2001], because of differences in resolution and horizontal correlation functions closer comparison is not warranted). Note that the forecast differences were at N144 resolution, that some of the tests reported in this paper used 3D-Var at N108 resolution (as does the operational 30-L system), but that the operational 40-L analyses and forecasts are at N48 resolution. The implied length scales at N108 (not shown) are shorter than those in Table 2, by more than 1/3 for the shortest scales, rather less for longer scales.

Figure 3 shows implied standard deviations of temperature calculated at N108. In the troposphere and lower stratosphere there are maxima at mid-latitudes (with slightly reduced values just round the tropopause) and generally small values in the tropics. In the stratosphere there is a very strong contrast between the dynamically active winter and the quiescent summer circulation (see, e.g., *Swinbank et al* [1998]). The sense of those seasonal differences are shown in these statistics, but the standard deviations are surprisingly large (particularly in the summer hemisphere), and may be a result of spurious gravity wave activity in the forecasts, as discussed later. Above about 5 hPa in the summer hemisphere values near the pole become particularly large. The temperature covariances implied at N48 are quite similar but with slightly smaller mid-latitude maxima. Both are slightly smoothed versions of the forecast difference covariances (not shown).

Figure 3 also shows the extent of linear (essentially geostrophic) balance in the forecast differences as measured by the variance of unbalanced pressure as a proportion of the total pressure variance. Below 100 hPa the extratropics are largely

balanced and the tropics are unbalanced – as expected. What is less expected is that above about 50 hPa the stratosphere is largely unbalanced at all latitudes (locally the balanced pressure can have the wrong sign giving ratios over 1); there is a hint of balance returning in the top few model levels – the lower mesosphere. In general the balance reaches higher in the dynamically active winter stratosphere than in the quieter summer stratosphere. Bouttier (2000, personal communication), and experience (by RS) with the NASA Data Assimilation Office (DAO) model confirms that stratospheric imbalance is not confined to the Met Office model. We will return to discussion of balance in the stratosphere in section 6.

4. Objective assessment of 3D-VAR

4.1 Impact of replacing AC with 3D-VAR in the 40-L configuration

Prior to implementation, trials of the 40-L Var assimilation system were run for two one-month periods (December 1998 and June 1999). Initial testing and comparisons were performed at a higher horizontal resolution (0.83° by 1.25°, N144) than was implemented operationally. This higher resolution is consistent with other operational trials, and was a first step towards extending the standard global model to span the stratosphere. Objective verification was performed using forecasts out to T+120, run from 12Z each day of each assimilation experiment. The forecasts were compared with observations and the 42-L AC analyses, which were the operational stratospheric products at the time, to give a set of statistics with which to evaluate the performance of each experiment.

As previously found for the standard (30-L) global model [*Lorenc et al*, 2000], the replacement of the AC scheme with 3D-Var led to a major improvement in forecast skill, as judged by verification against (primarily tropospheric) radiosonde observations. Figure 4 shows examples of comparisons of forecast scores for 100 hPa and 500 hPa height for the NH winter case, from forecasts out to 5 days. The 40-L AC trial results are shown by dotted lines and the 40-L Var trial results by full lines. There is a marked reduction in RMS Observation minus Forecast (O-F) errors over the entire forecast period in the new 40L-Var system. The same is also seen in virtually all of the winter hemisphere verification statistics. Figure 5 is an example of SH winter verification scores at 100 hPa, showing similar results. The summer hemisphere scores (Fig 5b) are also reduced, but less consistently and by a lesser amount. In the tropics (not illustrated), the results were not conclusive. A profile of NH winter RMS O-F height verification scores (Fig. 6) for T+120 shows the errors are reduced at all levels in the Var run.

Because of the sparse coverage of radiosonde observations in the stratosphere, verification was also performed against stratospheric analyses. The verification used the old operational 42-L AC analyses which has poorer tropospheric, but better stratospheric vertical resolution, and reduced (N48) horizontal resolution. Comparison of height verification scores shows the same improvement in the troposphere and lower stratosphere (e.g. at 100hPa, Fig. 7b) as was seen with the verification against observations. At higher levels (e.g. 10 hPa, Fig 7a), the AC and Var results become more similar. So, the 40-L Var trial is less consistently better than the 42-L AC in the stratosphere. But, considered overall, the geopotential height verification shows a clear improvement from the adoption of 3D-Var.

Figure 8 shows the verification of the forecast winds at 250 hPa. It shows the RMS O-F vector wind errors, for verification against radiosondes, calculated from the December tria for both NH and tropics. Again, the Var results show clear improvement to the scores over the entire forecast range. The equivalent plot for verification against analyses (not shown) is similar. There are some differences, because of the different way that the forecast results are sampled, but the impact of Var, compared to AC, is very similar.

Each trial was assessed using an index derived from the objective verification scores from selected fields, selected for their importance in operational forecasting. Table 3 shows the T+24 statistics, just a few of the many parameters used to compile this index operationally, in the 40-L Var and 40-L AC trials. It illustrates the positive impact of the Var change, shown in a reduction of the forecast RMS difference over virtually all fields.

4.2 Impact of extending configuration from 30-L to 40-L Var

The forecast results discussed in section 4.1 have also been compared with results from the 30-L Var system at the same horizontal resolution, in order to clarify the impact of the addition stratospheric levels. The 30-L Var results are plotted as dashed lines in the figures. In general, it was found that the tropospheric forecast skill of the 40-L Var system is very similar to that of the 30-L Var system of the same horizontal resolution. Figure 4 compares the geopotential height forecasts at both 500 hPa and 100 hPa in NH winter, and typifies the results found on most of the verifications against radiosondes. These results are consistent with the findings of *Lahoz* [1999], who found little benefit from additional levels on the skill of short range forecasts.

Figure 5 shows a marginal increase in 30-L Var skill over 40-L Var across a range of forecast times, but the SH results may need a little more care in their interpretation, especially higher up, as there is a lack of data here. What sometimes appears to be a slight degradation in the 40-L Var run may merely be due to the sampling of data, since there are far fewer observations with which to compare model data than the NH. In the SH, data are so sparse that some days have no observations at all at the highest levels, and therefore the results consist of a very few examples.

Profiles of forecast scores are shown in Fig. 6, for day 5 of the forecasts. As previously discussed, both the 30-L and 40-L Var show a considerable improvement over the 40-L AC results in the troposphere. Above 100 hPa, the impact of the improved stratospheric resolution in the 40-L model becomes much more noticeable. In all three experiments, there is a noticeable increase in the RMS forecast errors once one gets well into the stratosphere. But the increase is much steeper for the 30-L Var run, with very large errors at the highest verified level (10hPa). This increase in RMS error is a reflection of the very coarse model resolution; the top model layer in the 30-L model extends from 10 hPa to 0.5 hPa.

Calculation of the forecast skill index, in this case comparing 40-L Var over the 30-L Var trial results gave overall neutral results. A sample of the fields used appears in Table 4. Taking into account the 'tropospheric' bias of the fields used, and given that the two model formulations were almost identical in the troposphere, this

result showed what one might have expected; there were no significant differences resulting from the difference in error covariances.

5. Subjective comparisons of assimilation results in the stratosphere

In the stratosphere it harder to judge the impact of the change objectively, but subjective assessment indicates that the new system is generally performing well. In this section we focus on some examples taken from 10th November 2000, when the new 3D-VAR system was being tested in parallel to the old AC system, immediately prior to operational implementation.

Figs. 9 and 10 show examples of geopotential height, wind and temperature fields produced by both the new and old systems, plus similar results from the operational ECMWF data assimilation system. At (and below) 10 hPa (Figure 9) all three assimilation systems produce broadly similar results, particularly for height and wind. The temperature fields are somewhat more different. There is more small-scale structure in the ECMWF field, which seems unrealistic for the mid-stratosphere, but could be a consequence of the higher horizontal resolution in the ECMWF model. The centre of the polar vortex is about 2K colder in the new 3D-VAR results compared to either the old AC or the ECMWF results. The structure of the temperature maximum over eastern Siberia is more similar in the new 3D-VAR and ECMWF; it is possible that this similarity shows the benefit of assimilating radiances directly, rather than using retrievals.

At 1 hPa (Fig 10) the differences are more substantial, partly because the temperatures at this level are not well defined by the satellite soundings (all AMSU-A weighting functions peak below this level). Since the height fields represent the integrated effect of temperatures at lower levels, they are broadly similar. There are some major differences in the temperature fields. While all three generally show lower temperatures over the pole, there are large differences in the positioning and values of the minima. Both the 3D-VAR and ECMWF have a subsidiary temperature maximum over northern Russia, although they are different in detail. Additional changes between the AC and 3D-VAR systems are due to a correction of an error in the ozone climatology used by the old 42-L AC system; this probably accounts for the main differences at very large spatial scales, but not the positioning of maxima and minima.

Figure 11 compares the zonal-mean temperatures for the same day. The plots show zonal mean temperatures from the new 3D-VAR system, together with difference field (3D-VAR minus AC and 3D-VAR minus ECMWF). The (3D-VAR minus AC) temperature differences near the stratopause are dominated by the error in the ozone climatology in the old AC system. In addition, some differences will result from the poorer vertical resolution (and hence larger interpolation errors) in the new 40-level 3D-VAR system. After correction of the ozone climatology, the 3D-VAR results and ECMWF agree reasonably well, with differences of generally less than 2K at mid-stratosphere and below. The remaining discrepancies at upper levels are partly attributable to errors interpolating from relatively coarse model level near the stratopause, in both models.

Significant changes in low-latitude winds in the upper stratosphere (Figure 12) are attributed to insufficient balance between mass- and wind-field errors in the

background error covariance files in 3D-VAR. However, these differences are reduced in later months, so they are partly a reflection of slow spin-up of the upper stratospheric wind fields at low latitudes. Another factor is that the reduced stratospheric resolution in the 40-L 3D-VAR model (as compared to the 42-L AC model) is likely to degrade the representation of shear zones associated with the QBO and SAO, to some extent. Further work is planned with the aim of improving the analysis of stratospheric winds.

6. Discussion

6.1 Stratospheric Balance and the NMC method

The degradation of the 3D-VAR results, compared to AC, at lower latitudes is perhaps the most obvious symptom of the problems with imbalance in the stratosphere that was mentioned in section 3. One would anticipate that the atmospheric wind field would be at least as well balanced in the stratosphere as in the troposphere (with the possible exception of inertio-gravity waves), while results from the "NMC method" suggest the opposite.

At the current time, the NMC method is the only practical approach to the estimation of the climatological background error covariances for NWP. (An alternative approach developed by *Dee and De Silva* [1999] is restricted to regions well covered by radiosonde or similar data.) A version of the NMC method is currently used by most, if not all, modern data assimilation systems used for operational weather forecasting. However, the approach does have its limitations. It was originally introduced by *Parrish and Derber* [1992], following the observation that the differences between T+48 and T+24 forecast fields were a good proxy for the typical background errors in a T+6 forecast, although there is no firm theoretical justification for this linkage.

In the stratosphere-troposphere version of the AC scheme temperature increments were balanced by adding a fraction of the balancing geostrophic winds; this fraction was 0.7 in both the troposphere and stratosphere but with lower values round the tropopause [*Swinbank and O'Neill*, 1994]. In other words, a high degree of balance was imposed in the analysis increments.

By contrast, the degree of balance in the 3D-Var increments is derived entirely from the forecast statistics compiled using the NMC method. In the stratosphere, there are a range of gravity waves; generated by various processes, such as orographic drag, geostrophic adjustment and convection. For the most part, the waves are sub-gridscale. However, in numerical models, those processes tend to generate resolved waves which generally propagate up into the stratosphere. It is likely that these waves compromise the NMC method in the stratosphere.

It could be argued that, if the forecast fields contain spurious gravity waves, the imbalance in the background error statistics should allow observations of (realistically balanced) mass and wind fields to generate increments in such a way as to reduce the imbalance in the forecast. However, there are many mass field observations in the stratosphere and extremely few wind observations, so in practice it is impossible for the observations to correct any imbalance. When the analysis increments correct temperature errors (without balancing wind corrections), the

increments tend to be dispersed as gravity waves, rather than making associated changes to the wind fields (e.g. *Barwell and Bromley* [1988]). To some extent, this problem is alleviated by the IAU repeatedly inserting the increments, and thus limiting the generation of high frequency gravity waves, but this does not help the assimilation process very much where slower, inertia-gravity, waves are a problem.

6.2 Vertical Resolution

Another important factor is inadequate vertical resolution. Because of the higher static stability in the stratosphere vertical length scales will tend to be shorter there than in the troposphere (and mesosphere) – requiring higher vertical resolution there (e.g. *Lindzen and Fox-Rabinovitz* [1989]). For reasons of cost, the 40-L model has lower stratospheric vertical resolution than ideal; this almost certainly contributes to the lack of balance. The 30-L model, which has worse stratospheric resolution, becomes unbalanced at slightly lower levels. Vertical temperature correlations at high levels from the forecast differences become very narrow, reducing to delta functions when plotted as a function of model level. Under these circumstances we can't expect the differences to be balanced. (The unstaggered Lorenz grid doesn't help, either. The staggered Charney-Phillips vertical grid used in 3D-Var and being introduced with the New Dynamics should give better balance properties and eliminate a vertical interpolation.)

6.3 Effect of global linear balance

Several aspects of the global linear balance equation used can be questioned, some of these were tested numerically and results are summarised here. Use of gradient wind (non-linear balance) gave slightly better balance in the tropics and sub-tropics but little difference in mid-latitudes; the non-linear term will be larger in the stratosphere than the troposphere due to the higher wind speeds. *Randel* [1987] suggested that gradient winds were better than geostrophic winds in the stratosphere.

We calculate a balanced pressure field from the rotational wind, in the stratosphere with its long horizontal scales and short vertical ones it would probably be better to calculate the wind from the mass field. This would entail using a mass field analysis variable to represent the "balanced" component of the flow, complemented with "unbalanced" wind components (i.e. the reverse of what is used the Met Office 3D-Var system); this was not tried. We might well get better results in the stratosphere if the "balanced" analysis variable were more closely related to our primary set of observations.

Application of the global linear balance equation level by level gives insufficient vertical consistency of the balanced mass field; in particular the implied temperature covariances can be unrealistic. A vertical ridge regression (p 3011 of *Lorenc et al* [2000]) is applied to overcome this. Ordinary regression gives more balance in the winter stratosphere but very noisy regression coefficients – apparently overfitting the data. Various 30-L tests were run using modified regression methods but results were mixed and overall slightly worse although there were some improvements in the stratosphere.

6.4 Possible improvements

In an effort to alleviate the problems just discussed, some simple adaptations of the NMC method have already been tried. Time filtering the forecast fields over six-hours (instead of using instantaneous T+24 and T+48 forecasts) gave a slight reduction in the proportion of unbalanced pressure. Applying divergence damping during the forecasts increased rather than reduced A_p. However it did reduce the variance of the divergent winds in the stratospheric forecast differences – this variance seems rather large. So, neither of these filtering approaches appeared to give sufficient benefit to be worthwhile pursuing.

A direct way of reducing the imbalance in the background error covariances is an artificial reduction of the A_p error variance (while retaining the full variance of ψ, and possibly also χ). This has not yet been attempted, and care would need to be taken not to make the A_p variance too small. At the DAO, a similar decision was taken to artificially increase the degree of balance in the stratospheric increments. While a different approach was used (because of major differences in formulation of their assimilation system), the result was an overall improvement in forecast skill.

Another likely problem area is the use of a mass-weighting $\mathbf{P}_v$ in the calculation of vertical modes (Eq. 6). While this does not directly effect balance in the error statistics, it gives the tropospheric forecast errors (inferred from forecast differences) much higher statistical weight than the stratospheric errors, where the air density is much lower. This is likely to reduce the robustness of the background error statistics in the stratosphere (although it might improve the robustness of the statistics at low levels). It is likely that uncertainities in the statistical calculation have reduced the degree of balance in the stratosphere, and possibly also contributed to the large stratospheric error variances illustrated by Fig 3. To address this issue, we are planning to experiment with the use an alternative formulation of $\mathbf{P}_v$ based on the (altitude) thickness of the model layers; this approach has been used by the US Naval Research Laboratory (NRL) [*Daley and Barker*, 2000]. We hope this will lead to improved reliability in the stratospheric error statistics, without significant degradation of the tropospheric error statistics.

7. Conclusions and Future plans

The results from the operational trials presented in this paper show that, for the most part, the stratosphere-troposphere version of the 3DVAR system performs well. The objective forecast verification scores show the benefit of replacing the former AC scheme with 3DVAR. The scores also show that the extra stratospheric levels have had a positive impact near the tropopause and at higher altitudes. Subjective assessment of the stratospheric analyses shows that 3DVAR generally produces slightly smoother, and probably more realistic meteorological fields – although that might also be an effect of the reduced vertical resolution.

There are some systematic changes in the upper stratosphere, which were partly a result of correcting a long-standing error in the ozone climatology. The improvement in the temperature structure has brought the Met Office analyses closer to the ECMWF results. However, at lower latitudes there has been a degradation in the 3DVAR results, as compared to the AC results. This is probably a reflection of the inadequate balance in the stratosphere, as discussed in section 6.

The next step in the development of the stratosphere-troposphere data assimilation system will be the adoption of the new semi-implicit dynamical core which has been developed for the Unified Model. The current stratospheric 3DVAR system is being ported to the new version of the UM with minimal changes. At the time of writing, that work is well advanced. It is planned to implement that upgrade in the operational suite during autumn 2002. In parallel with those developments, a version of 3D-VAR is being developed to assimilate ozone observations; that work is being carried out in collaboration with the NERC Data Assimilation Research Centre (DARC).

In the next couple of years, it is planned that the stratosphere-troposphere data assimilation and forecast system will be used to demonstrate the feasibility of extending the standard operational global forecast model to span the stratosphere. As a step towards that goal, further work will be done to try and alleviate the problems with stratospheric balance, following the ideas discussed in section 6.4. In addition, work will be done on a revised bias correction scheme for satellite radiances. Once this work is completed, we will be able to satisfy all the customer requirements for global forecasts, using a single model configuration.

Acknowledgements We would like to thank many people in NWP for helping us to set up and run the stratospheric data assimilation trials, particularly Andrew Smith, Colin Parrett, Dave Matthews and Jim Bolton. We thank Adam Scaife and David Jackson for advice regarding the stratospheric configuration of the UM, and Adam Scaife, Andy White, Steve Cohn and Mike Cullen for discussions on stratospheric balance.

References

Barwell, B. and R. Bromley, 1988: The adjustment of numerical weather prediction models to local perturbations, *Quart. J. R. Meteorol. Soc.*, **114**, 665-684.

Bloom, S.C., L.L.Takacs, A.M. da Silva and D. Ledvina, 1996: Data assimilation using incremental analysis updates, *Mon Weather Rev.*, **124**, 1256-1271.

Boer, G.J. 1983: Homogeneous and isotropic turbulence on the sphere. *J. Atmos. Sci.*, **40,** 154-163.

Charney, J.G and P.G. Drazin, 1961: Propagation of planetary-scale disturbances from the lower into the upper atmosphere. *J. Geophys. Res.*, **66**, 83-109.

Cohn, S.E., A. da Silva, J. Guo, M. Sienkiewicz and D. Lamich, 1998: Assessing the effects of data selection with the DAO Physical-space Statistical Analysis System, *Mon. Weather Rev.*, **126**, 2912-2926.

Courtier, P., J.N. Thepaut and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach, *Quart. J. R. Meteorol. Soc.*, **120**, 1389-1408.

Courtier P, E. Andersson, W. Heckley, J. Pailleux, D. Vasiljevic, M. Hamrud, A. Hollingsworth, F. Rabier and M. Fisher, 1998: The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: Formulation . *Quart. J. R. Meteorol. Soc.*, **124,** 1783-1807.

Cox, P.M., R.A. Betts, C.B. Bunton, R.L.H. Essery, P. R. Rowntree and J. Smith, 1999: The impact of new land surface physics on the GCM simulation and climate and climate sensitivity. *Climate Dynamics*, **15**, 183-203.

Cullen, M.J.P., 1993: The unified forecast / climate model, *Meteorol. Mag.*, **122**, 81-94.

Daley, R. and E. Barker, 2000: The NAVDAS sourcebook. Naval Research Laboratory NRL/PU/7530-00-418, 151pp.

Dee, D.P. and A. da Silva, 1999: Maximum-likelihood estimation of forecast and observation error covariance parameters. Part I: Methodology, *Mon. Wea. Rev.*, **127**, 1822-1834.

Derber, J. and F. Bouttier, 1999: A reformulation of the background error covariances in the ECMWF global data assimilation system. *Tellus,* **51A,** 195-221.

Edwards, J.M. and A. Slingo, 1996: Studies with a flexible new radiation code. I: choosing a configuration for a large scale model. *Quart. J. Roy. Meteor. Soc.*, **122**, 689-719.

English, S.J., R.J. Renshaw, P.C. Dibben, A.J. Smith, P.J. Rayer, C. Poulsen, F.W. Saunders and J.R. Eyre, 2000: A comparison of the impact of TOVS and ATOVS satellite sounding data on the accuracy of numerical weather forecasts. *Quart. J. R. Meteorol. Soc.*, **126**, 2911-2931

Gaspari, G. and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Met. Soc.*, **125**, 723-758

Harris, B.A. and G. Kelly 2001: A satellite radiance-bias correction scheme for data assimilation. *Quart. J. R. Meteorol. Soc.*, **127**, 1453-1468.

Ide, K., P. Courtier, M. Ghil, and A.C. Lorenc, 1997: Unified notation for data assimilation: Operational, sequential and variational. *J. Met. Soc. Japan*, **75**, 181-189.

Ingleby, N.B., 2001: The statistical structure of forecast errors and its representation in The Met.Office global 3-D Variational Data Assimilation Scheme. *Quart. J. R. Meteorol. Soc.*, **127**, 209-231.

Ingleby, N.B., 2002: Estimation of Background Error Statistics. *VSDP 20,* available from NWP Division, Met Office.

Ingleby, N.B., S P Ballard, R S Bell, A Berney, B Candy, S English, D Li, A C Lorenc, R J Renshaw and J K Ridley, 2001: Recent developments of the Met Office global 3D-Var system. *preprints AMS 18th Conference on WAF/14th Conference on NWP*, 1-4,

Jackson, D.R., J. Austin and N. Butchart, 2001: An updated climatology of the troposphere-stratosphere configuration and the Met Office's Unified Model. *J. Atmos. Sci.*, **58**, 2000-2008.

Lahoz, W.A, 1999: Predictive skill of the UKMO Unified Model in the lower stratosphere. *Quart. J. R. Meteorol. Soc.*, **125**, 2205-2238

Lindzen, R.S. and M. Fox-Rabinovitz, 1989: Consistent vertical and horizontal resolution. *Mon. Wea. Rev.*, **117**, 2575-2583.

Lorenc, A.C., R.S. Bell and B. Macpherson, 1991: The Meteorological Office Analysis Correction Data Assimilation scheme, *Quart. J. R. Meteorol. Soc.*, **117**, 59-89.

Lorenc, A.C., S.P. Ballard, R.S. Bell, N.B. Ingleby, P.L.F. Andrews, D.M. Barker, J.R. Bray, A.M. Clayton, T. Dalby, D. Li, T.J. Payne and F.W. Saunders, 2000: The Met. Office global 3-dimensional variational data assimilation scheme, *Quart. J. R. Meteorol. Soc.*, **126**, 2991-3012.

Malcolm, A.J., T. Davies, S. Bell and A. Clayton, 2001: Data assimilation and forecast tests of a new integration scheme for the Met Office Unified Model, *Proceedings of 18th conference of Weather Analysis and Forecasting and 14th conference on Numerical Weather Prediction*, J99-J102, American Meteorological Society.

Milton, S.E. and C.A. Wilson, 1996: The impact of parametrized subgrid-scale orographic forcing on systematic errors in a global NWP model. *Mon. Wea. Rev.*, **124**, 2023-2045.

Palmer, T.N., G.J. Shutts and R. Swinbank, 1986: Alleviation of a systematic westerly bias in general circulation and numerical weather prediction models through an orographic gravity wave drag parametrization, *Quart. J. R. Met. Soc.*, **112,** 1001-1039.

Parrish, D.F. and J.C. Derber, 1992: The National Meteorological Center's spectral statistical interpolation analysis scheme. *Mon. Wea. Rev.*, **120,** 1747-1763.

Randel, W.J. 1987: The evaluation of winds from geopotential height data in the stratosphere. *J. Atmos. Sci.*, **44,** 3097-3120.

Richman, M.B., 1986: Rotation of principal components. *J. Climatol.*, **6**, 293-335.

Swinbank, R. and A. O'Neill, 1994: A stratosphere-troposphere data assimilation system, *Monthly Weather Review,* **122**, 686-702.

Swinbank, R., W.A. Lahoz, C.S. Douglas, A. Heaps, A. O'Neill and D.Podd, 1998: Middle Atmosphere variability in the Met Office Unified Model. *Quart. J. Roy. Met. Soc.*, **124**, 1485-1525.

Table 1. Differential length-scales (km) for forecast differences, selected model levels (cf. Table 1 of Ingleby, 2001). The second column gives the average pressures in hPa for the levels. Due to the vertical staggering T and RH are at slightly lower average pressures than those indicated. See text for definition of symbols.

Level	P	ψ	RKE	χ	DKE	A_p	RH	P	T
34	3.9	1287	470	1547	357	797		928	444
29	26	1148	430	1511	352	912		913	537
25	102	1131	353	1368	317	682	303	764	438
19	251	747	332	1030	261	541	250	599	342
11	867	683	286	884	223	491	201	546	334
4	867	711	271	1100	231	498	197	534	297
2	961	634	246	746	212	428	180	500	242

Table 2. Implied differential length-scales (km) for N48 resolution, selected model levels (cf. Table 1).

Level	P	ψ	RKE	χ	DKE	A_p	RH
34	3.9	729	363	760	347	608	
29	26	697	356	834	364	598	
25	102	649	335	761	347	521	318
19	251	602	338	688	337	470	329
11	519	595	335	634	317	491	331
4	867	592	329	655	325	495	328
2	961	577	322	616	316	469	324

Table 3: Scores from selected fields against observations of 40l-Var over 40l-AC used to compile the 'NWP index' used to assess performance of trial runs for each of the principle T+24 fields (many more in full index).

Area	Field	FC Range	40L-AC FC RMS	40L-Var FC RMS	FC RMS Diff (%)
NH	PMSL	T+24	2.477	2.076	-16.18
NH	H500	T+24	21.992	18.778	-14.62
NH	W250	T+24	8.016	7.293	-9.01
Trop	W850	T+24	4.422	4.299	-2.78
Trop	W250	T+24	6.991	6.409	-8.33
SH	PMSL	T+24	2.038	1.868	-8.36
SH	H500	T+24	21.120	19.781	-6.34
SH	W250	T+24	8.552	7.999	-6.47

Table 4: Scores from selected fields against observations of 40l-Var over 30L-Var used to compile the 'NWP index' used to assess performance of trial runs for each of the principle T+24 fields (many more in full index).

Area	Field	FC Range	30L-Var FC RMS	40L-Var FC RMS	FC RMS Diff (%)
NH	PMSL	T+24	2.093	2.076	-0.80
NH	H500	T+24	18.700	18.778	0.41
NH	W250	T+24	7.386	7.293	-1.25
Trop	W850	T+24	4.272	4.299	0.63
Trop	W250	T+24	6.426	6.409	-0.26
SH	PMSL	T+24	1.886	1.868	-0.95
SH	H500	T+24	19.112	19.781	3.50
SH	W250	T+24	8.034	7.999	-0.44

Figure Captions

Figure 1. Comparison of three different configurations of the Unified Model (lines indicate model layer boundaries): a) former (42-L) stratospheric data assimilation model; b) new 40-L stratospheric data assimilation model; c) 30-L operational global model.

Figure 2: Vertical modes for rotational kinetic energy

Figure 3: (a) The ratio of the variance of unbalanced pressure to the total pressure variance, showing the extent of linear (essentially geostrophic) balance in the forecast differences. (b) Implied standard deviations of temperature calculated at N108

Figure 4: Height verified against radiosonde observations, RMS O-F for December trial, NH winter: full line 40-L VAR, dashed line 30-L VAR and dotted line 40-L AC (a) 100 hPa (b) 500 hPa.

Figure 5: 100 hPa height verified against radiosonde observations, RMS O-F for June trial, NH winter: full line 40-L VAR, dashed line 30-L VAR and dotted line 40-L AC (a) SH winter (b) NH summer.

Figure 6: Profile of RMS forecast scores height verified against observations at T+120 for Dec trial, NH winter: full line 40-L VAR, dashed line 30-L VAR and dotted line 40-L AC.

Figure 7: Height forecast scores verified against 42-L AC analyses for NH winter: full line 40-L VAR, dashed line 30-L VAR and dotted line 40-L AC (a) 10 hPa (b) 100 hPa

Figure 8: 250 hPa Wind RMS vector error verified against Sonde observations for December: full line 40-L VAR, dashed line 30-L VAR and dotted line 40-L AC (a) NH (b) Tropics

Figure 9. 10 hPa geopotential height, wind (upper) and temperature (lower) fields valid at 12 UTC on 10 November 2000, produced by the 42-L AC (left), operational ECMWF (centre) and new 40-L 3D-VAR (right) assimilation systems.

Figure 10. As Fig. 9, but for 1 hPa.

Figure 11. Comparison of zonal-mean temperatures for 10th November 2000. The plots show results from the new 40-L 3D-VAR (top), 3D-VAR minus ECMWF (centre) and 3D-VAR minus 42-L AC (bottom).

Figure 12. As Fig 11, but for zonal-mean zonal winds.

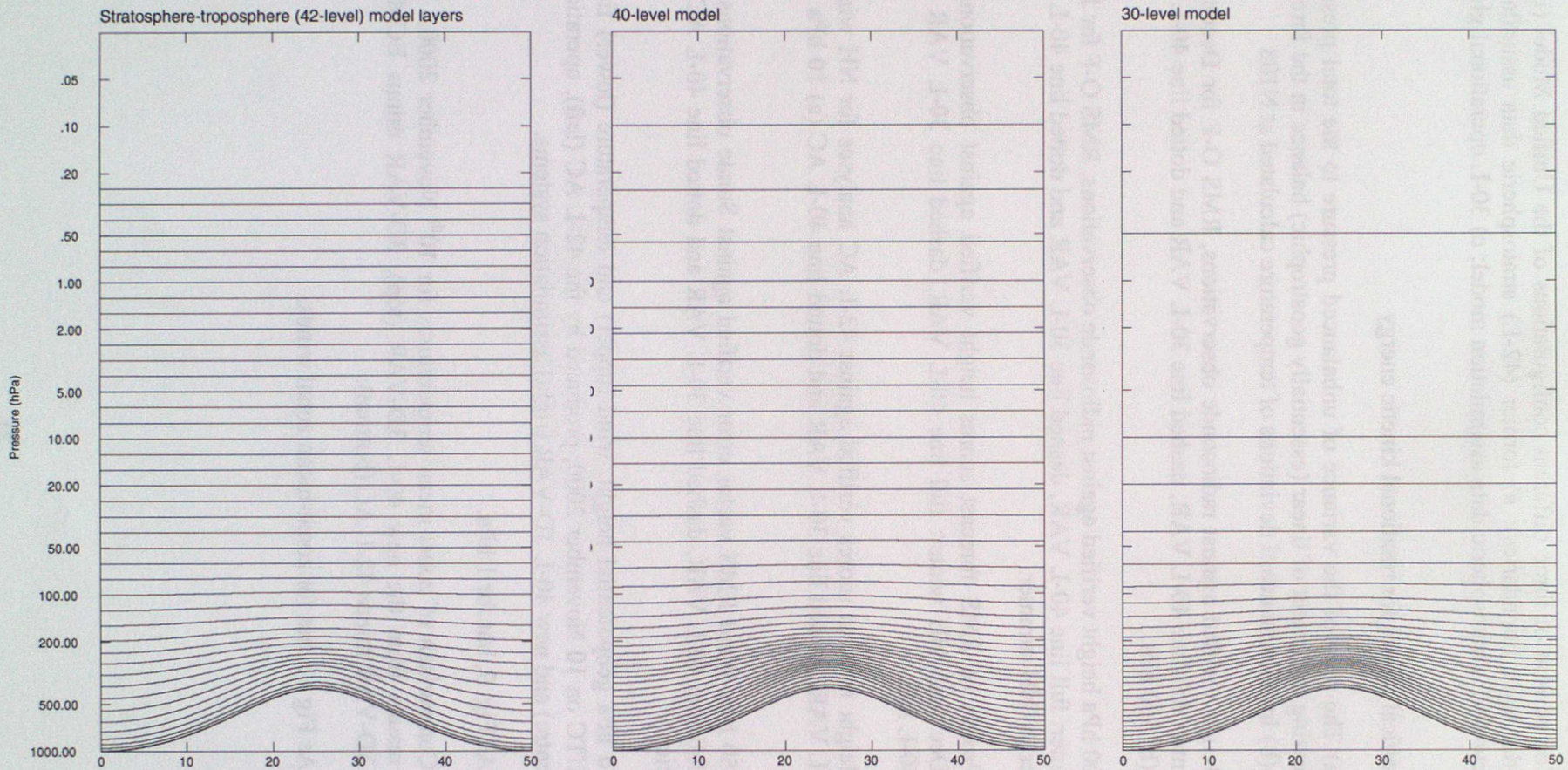


Figure 1. Three sets of levels for the Met Office Unified Model: a) former 42-level stratospheric data assimilation system; b) new 40-level 3D-VAR stratospheric data assimilation system and c) current 30-level standard global assimilation system

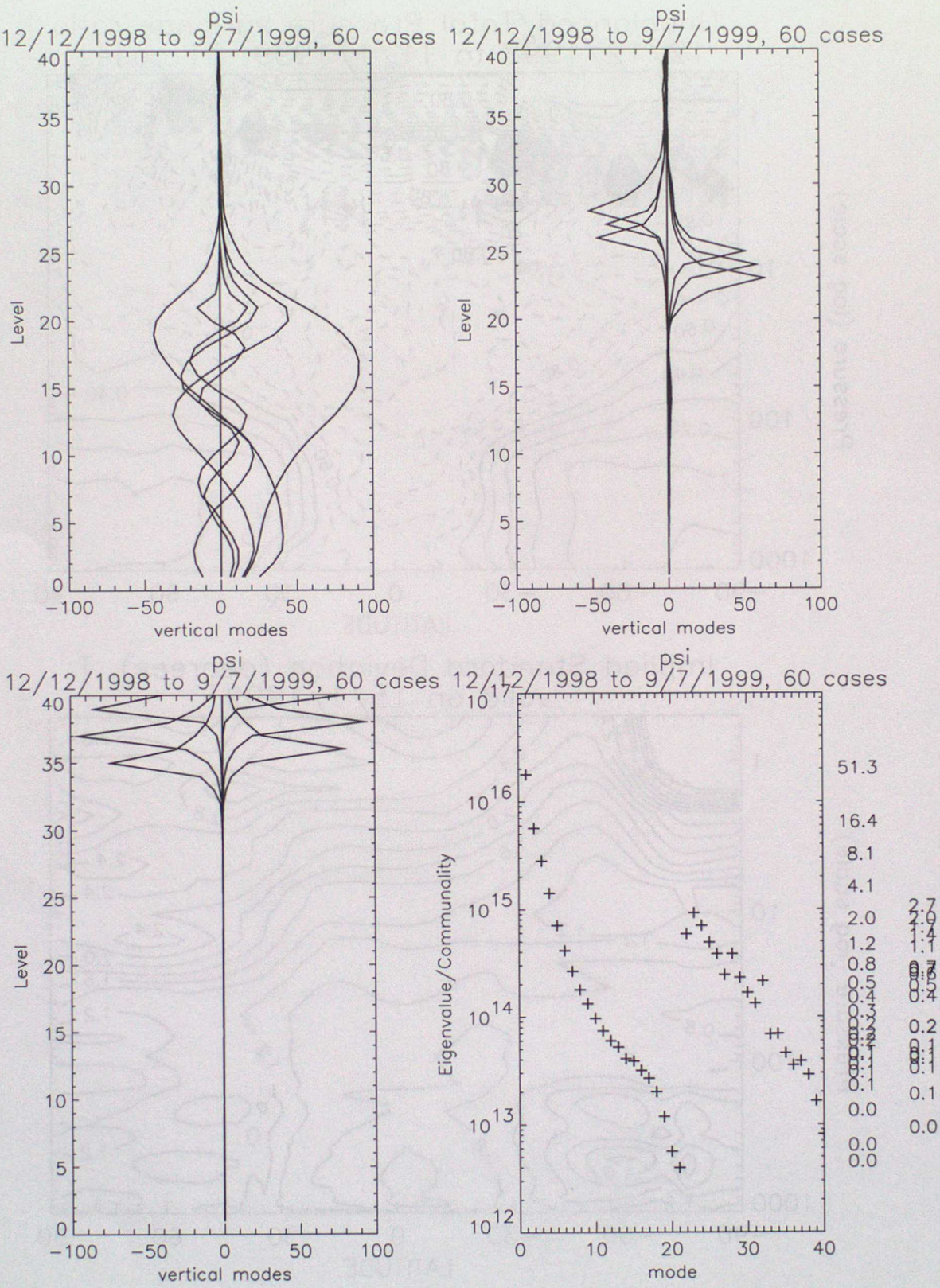


Figure 2. Vertical modes for rotational kinetic energy

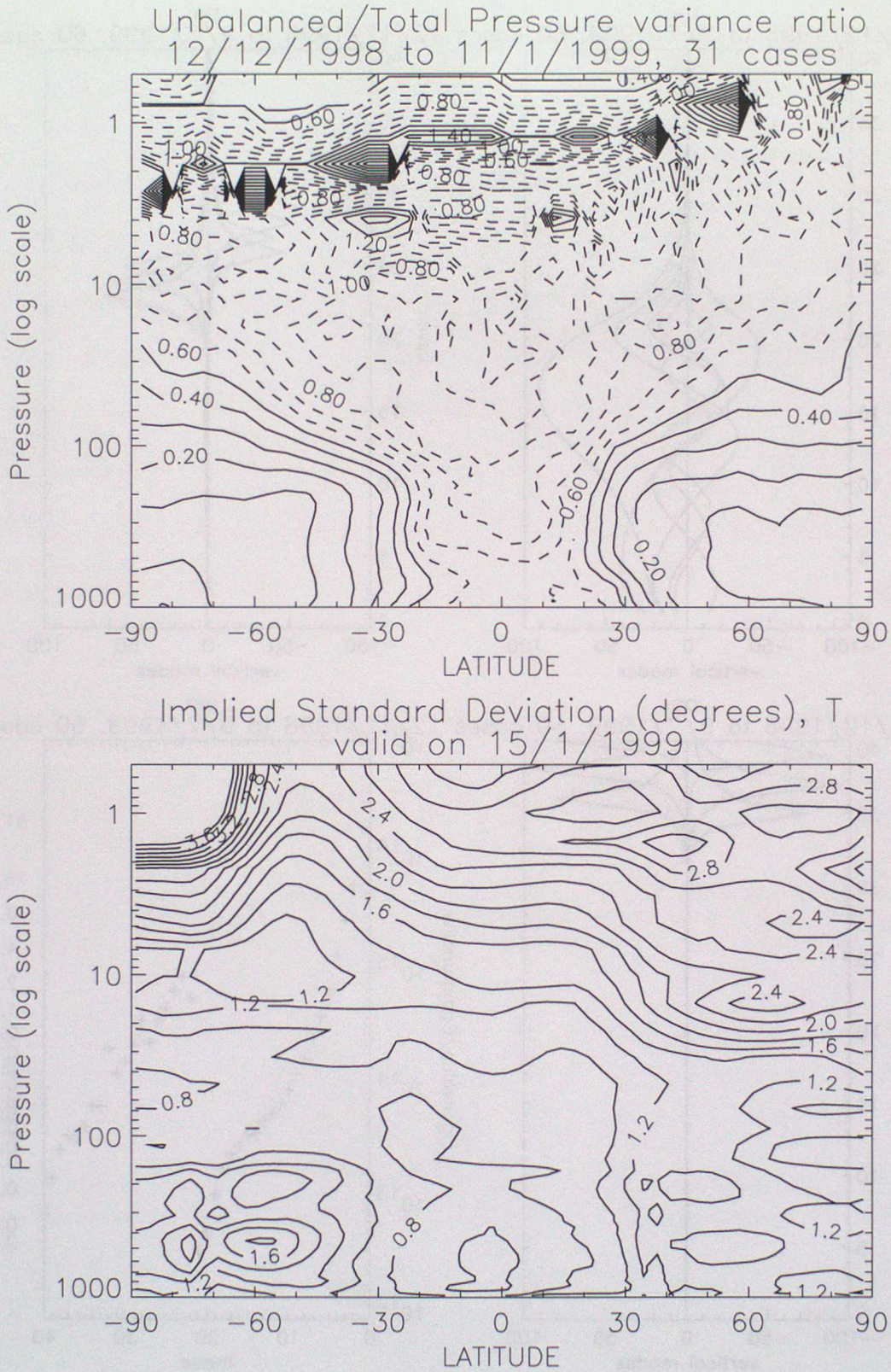


Figure 3. (a) The ratio of the variance of unbalanced pressure to the total pressure variance, showing the extent of linear (essentially geostrophic) balance in the forecast differences. (b) Implied standard deviations of temperature calculated at N108.

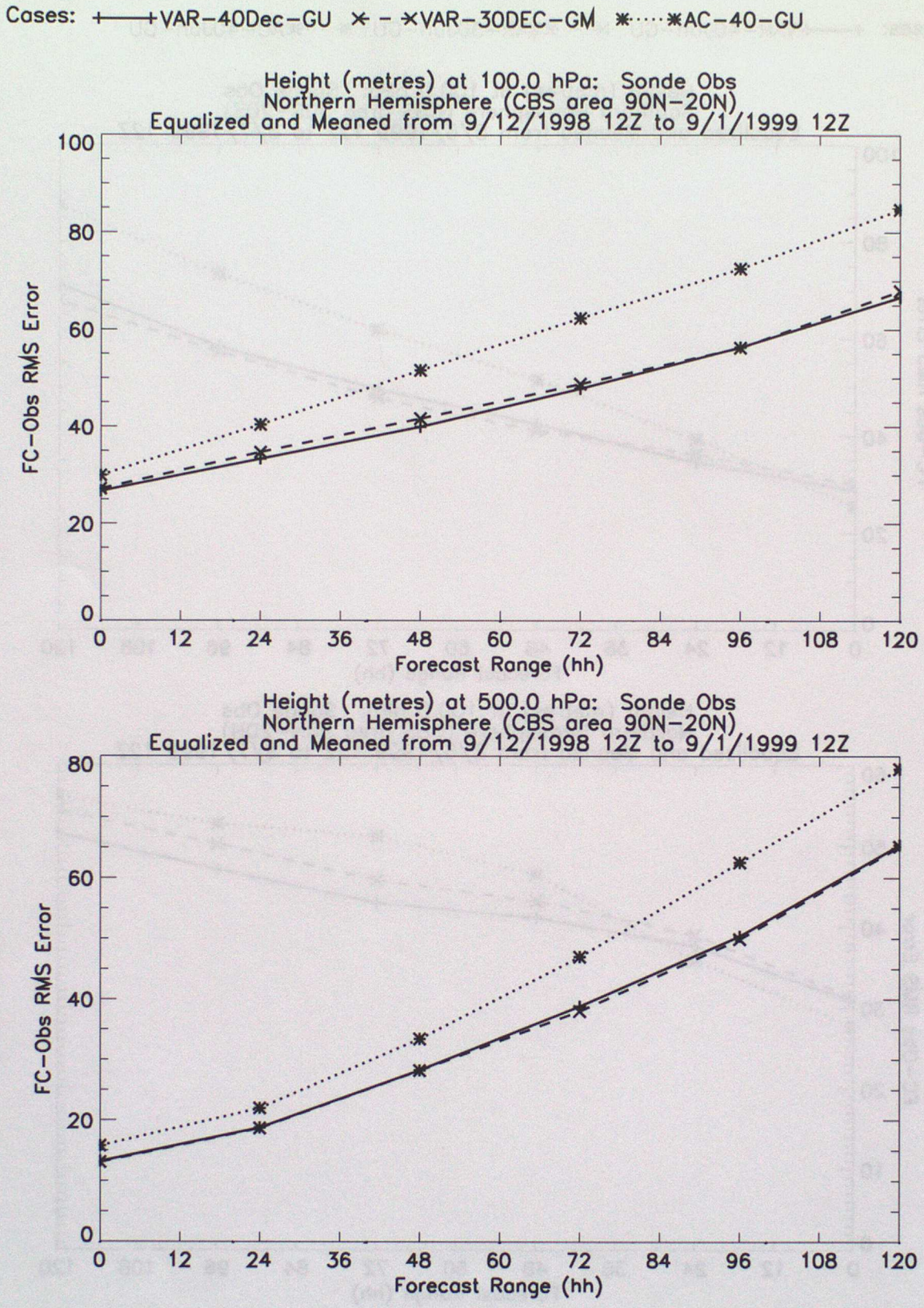


Figure 4: Height verified against radiosonde observations, RMS O-F for December trial, NH winter: full line 40-L VAR, dashed line 30-L VAR and dotted line 40-L AC (a) 100 hPa (b) 500 hPa.

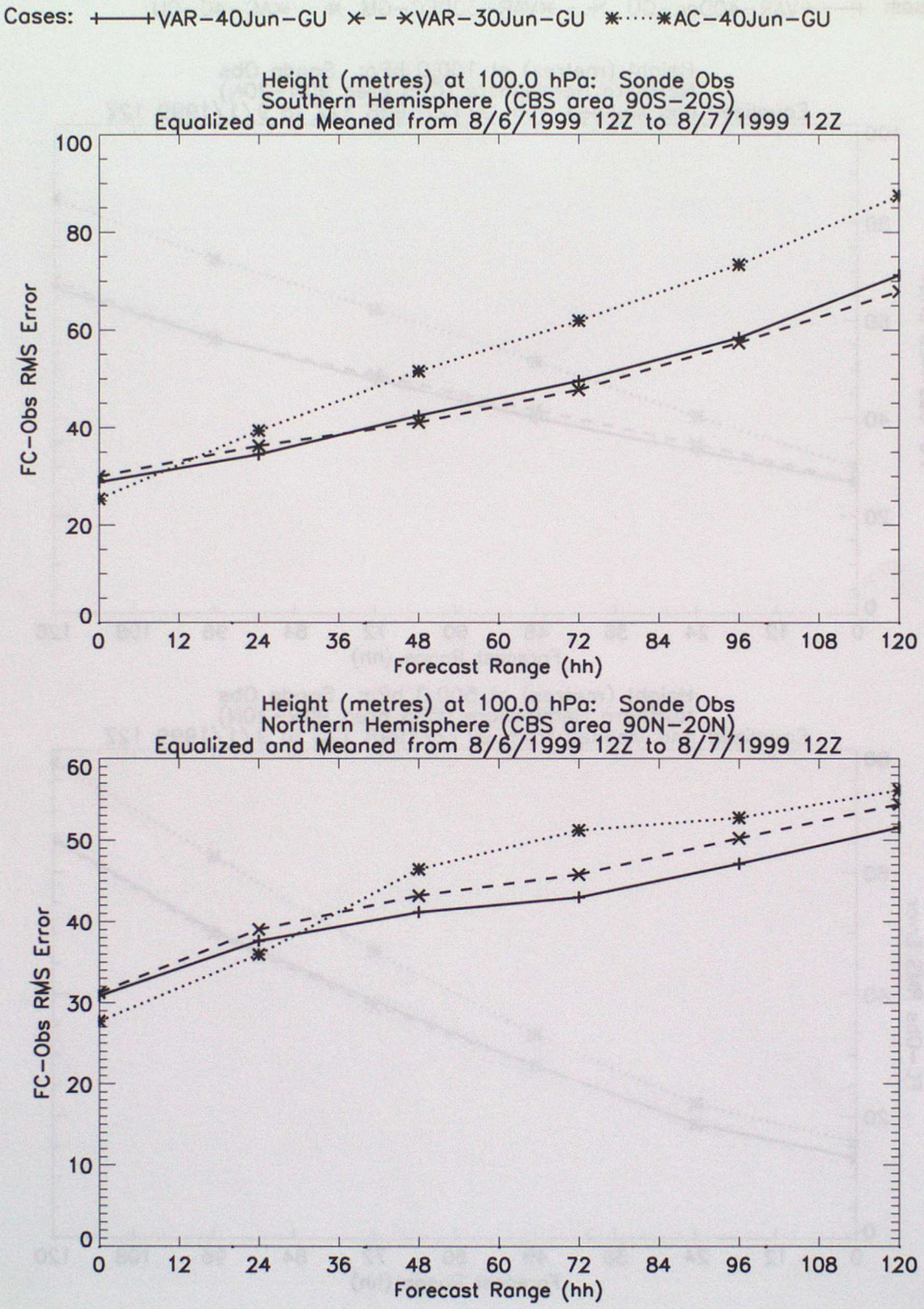


Figure 5: 100 hPa height verified against radiosonde observations, RMS O-F for June trial, NH winter: full line 40-L VAR, dashed line 30-L VAR and dotted line 40-L AC (a) SH winter (b) NH summer.

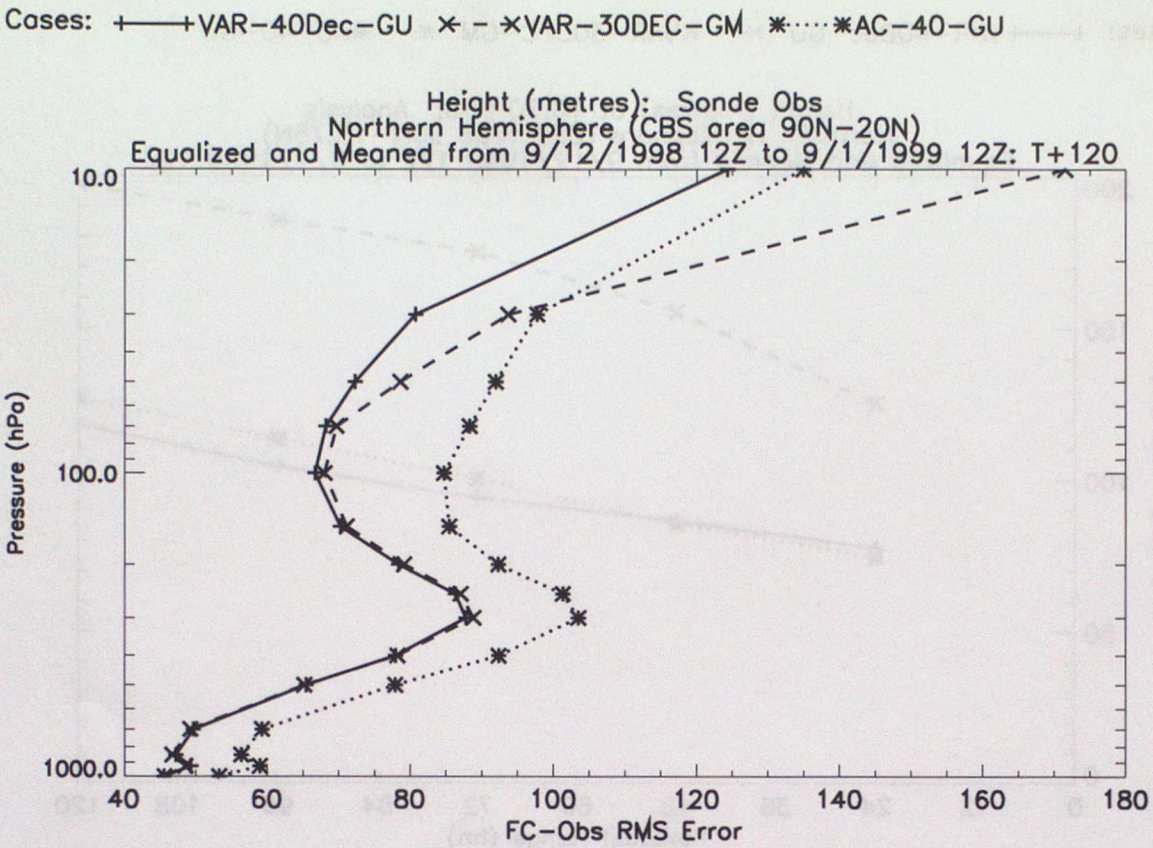


Figure 6: Profile of RMS forecast scores height verified against observations at T+120 for Dec trial, NH winter: full line 40-L VAR, dashed line 30-L VAR and dotted line 40-L AC.

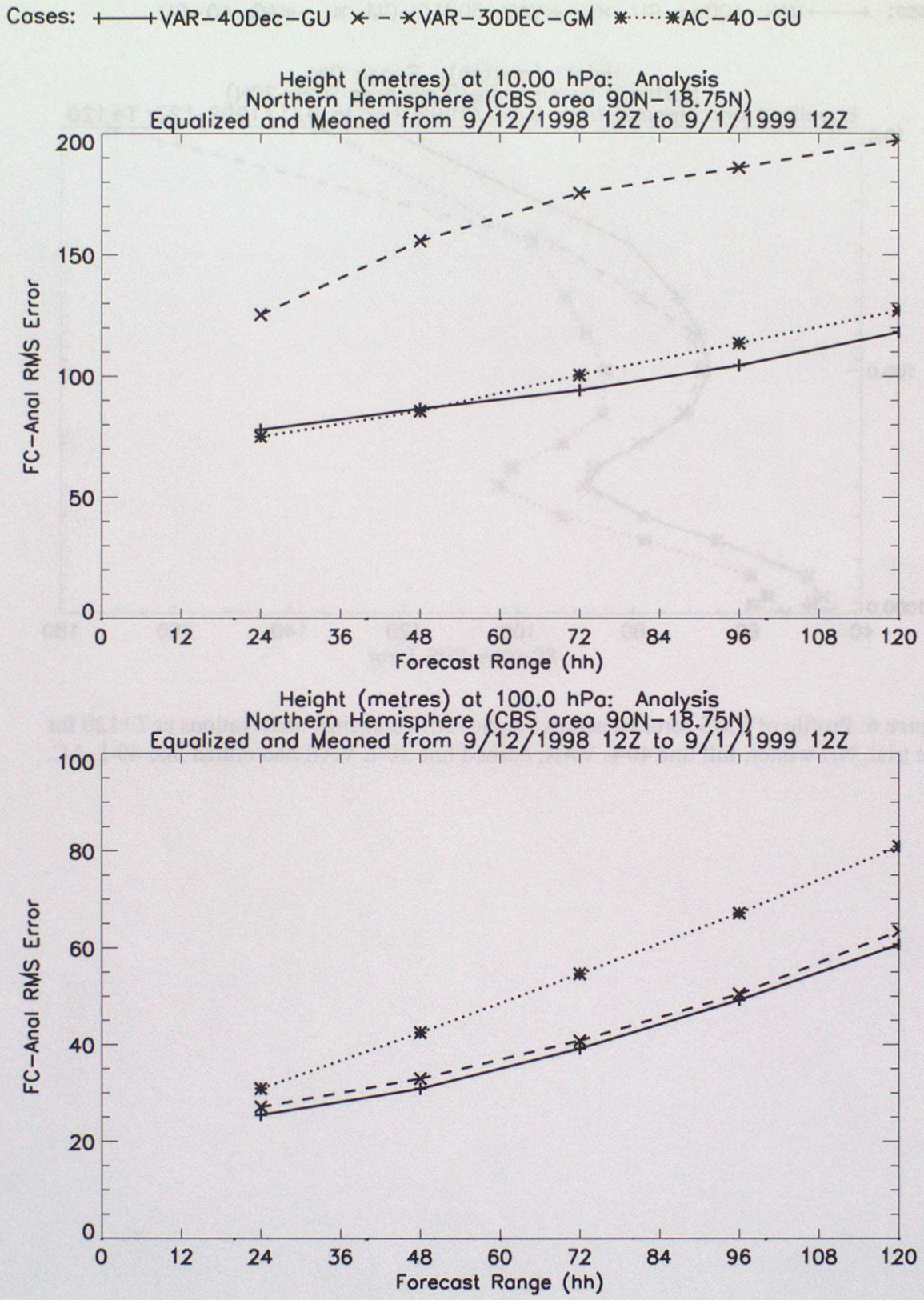


Figure 7: Height forecast scores verified against 42-L AC analyses for NH winter: full line 40-L VAR, dashed line 30-L VAR and dotted line 40-L AC (a) 10 hPa (b) 100 hPa

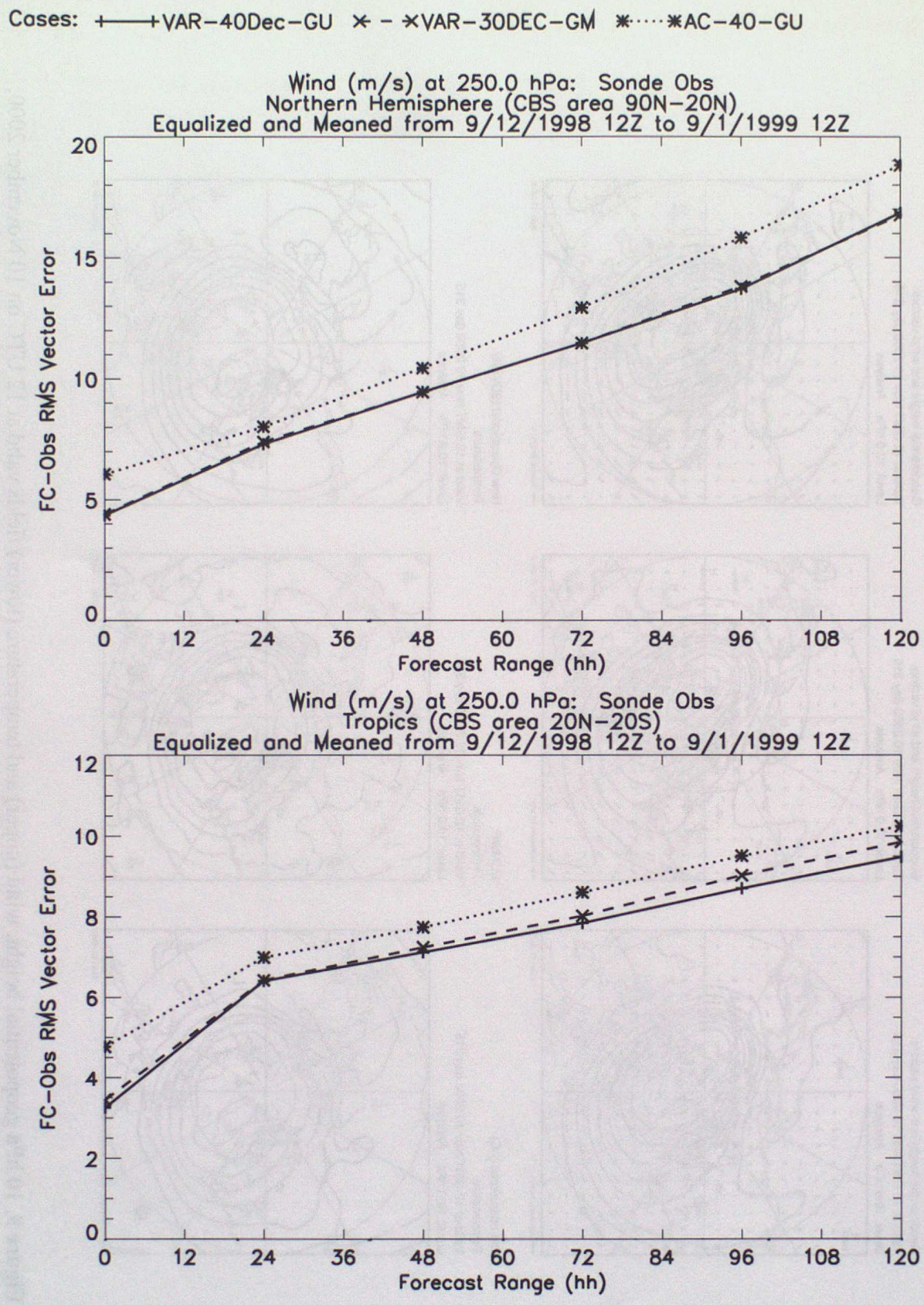


Figure 8: 250 hPa Wind RMS vector error verified against Sonde observations for December: full line 40-L VAR, dashed line 30-L VAR and dotted line 40-L AC (a) NH (b) Tropics

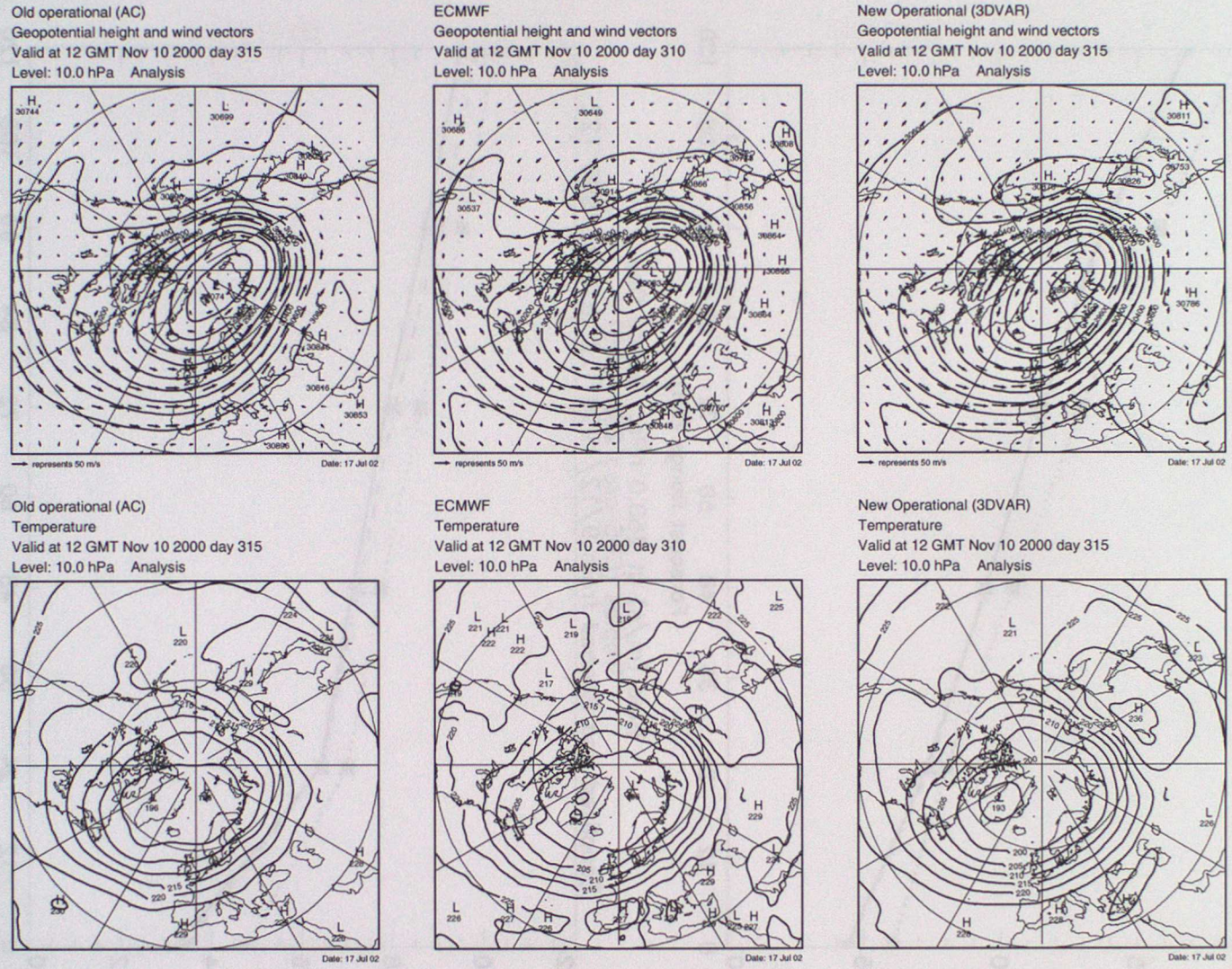


Figure 9. 10 hPa geopotential height, wind (upper) and temperature (lower) fields valid at 12 UTC on 10 November 2000, produced by the 42-L AC (left), operational ECMWF (centre) and new 40-L 3D-VAR (right) assimilation systems.

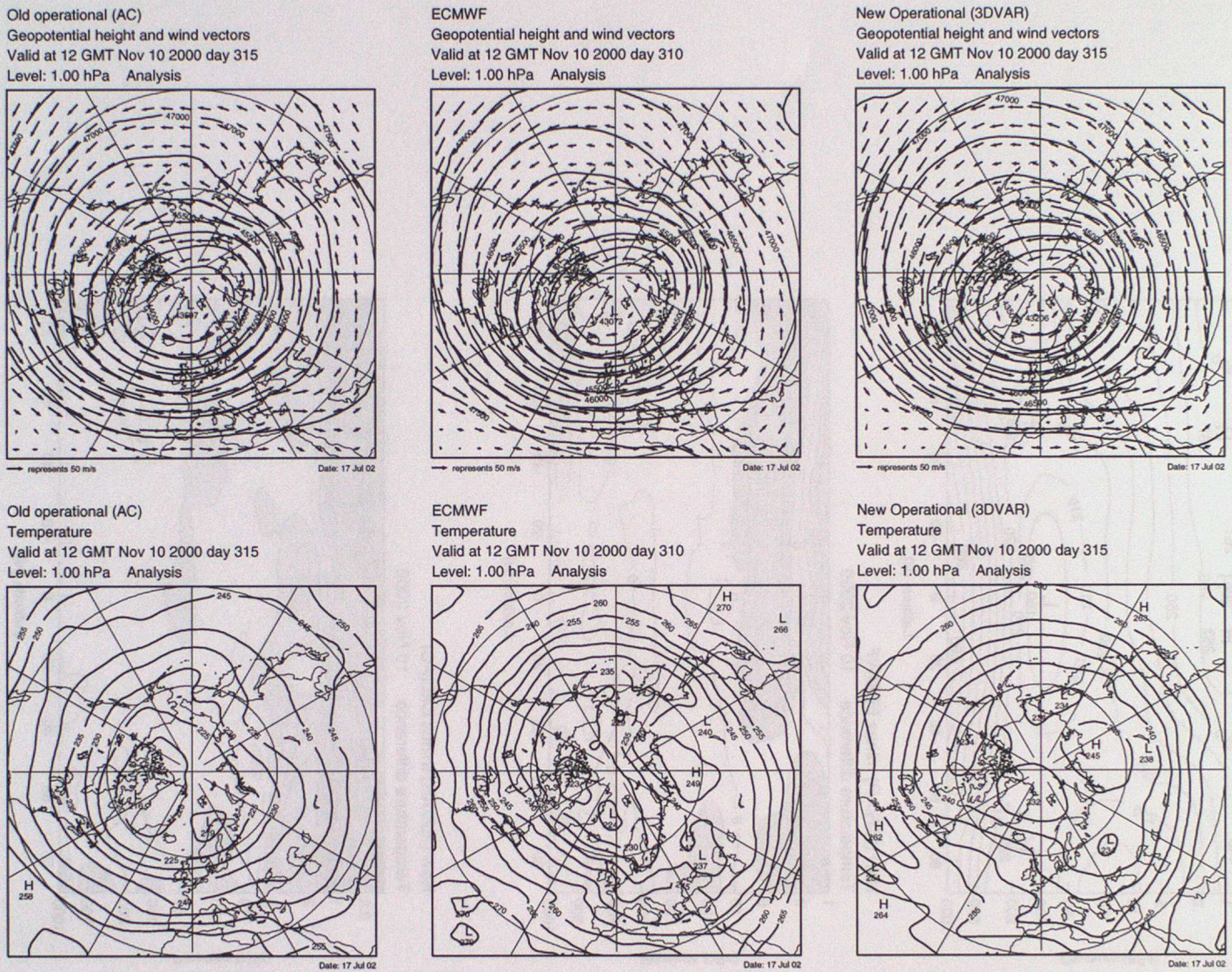


Figure 10. As Fig. 9, but for 1 hPa.

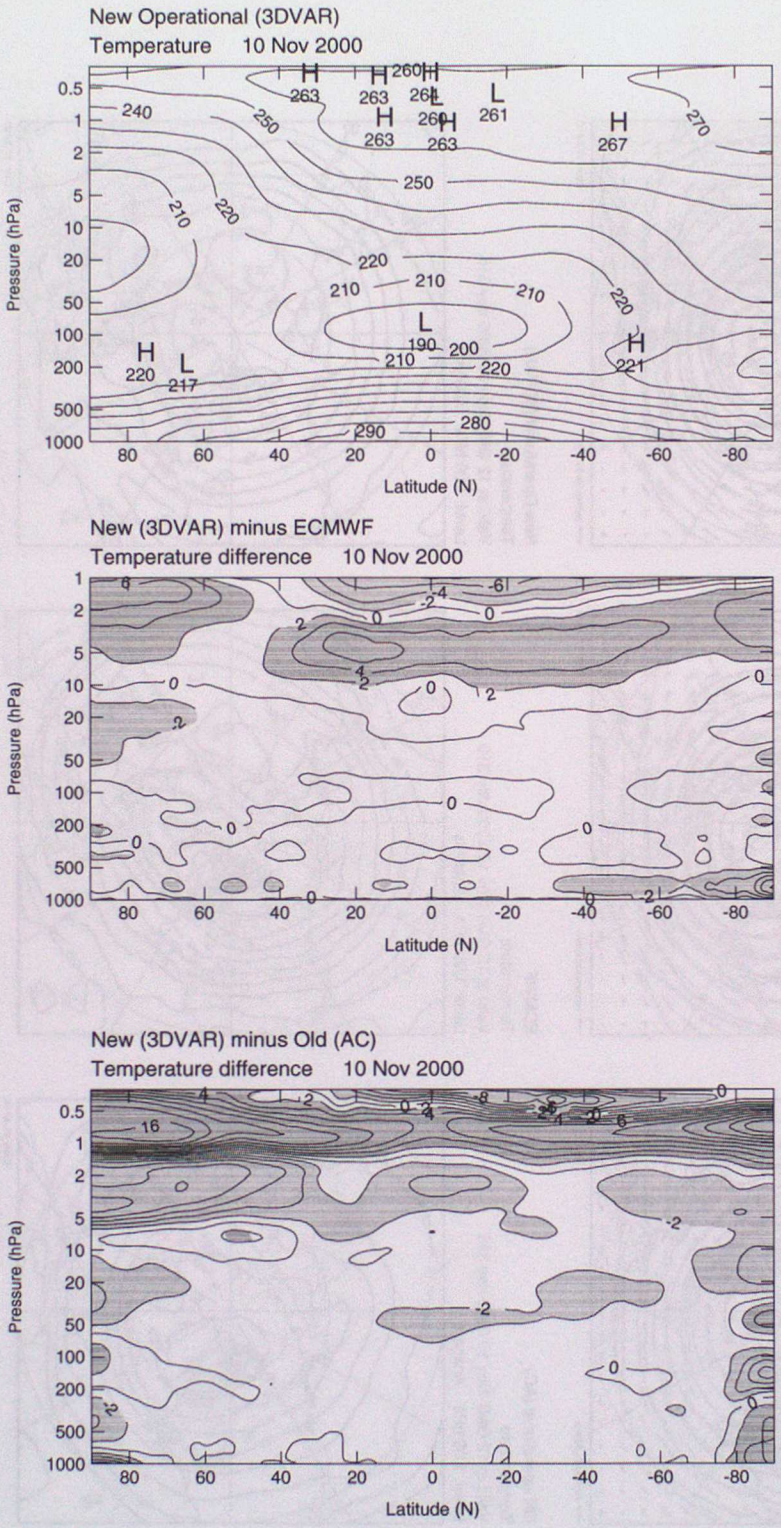


Figure 11. Comparison of zonal-mean temperatures for 10th November 2000. The plots show results from the new 40-L 3D-VAR (top), 3D-VAR minus ECMWF (centre) and 3D-VAR minus 42-L AC (bottom).

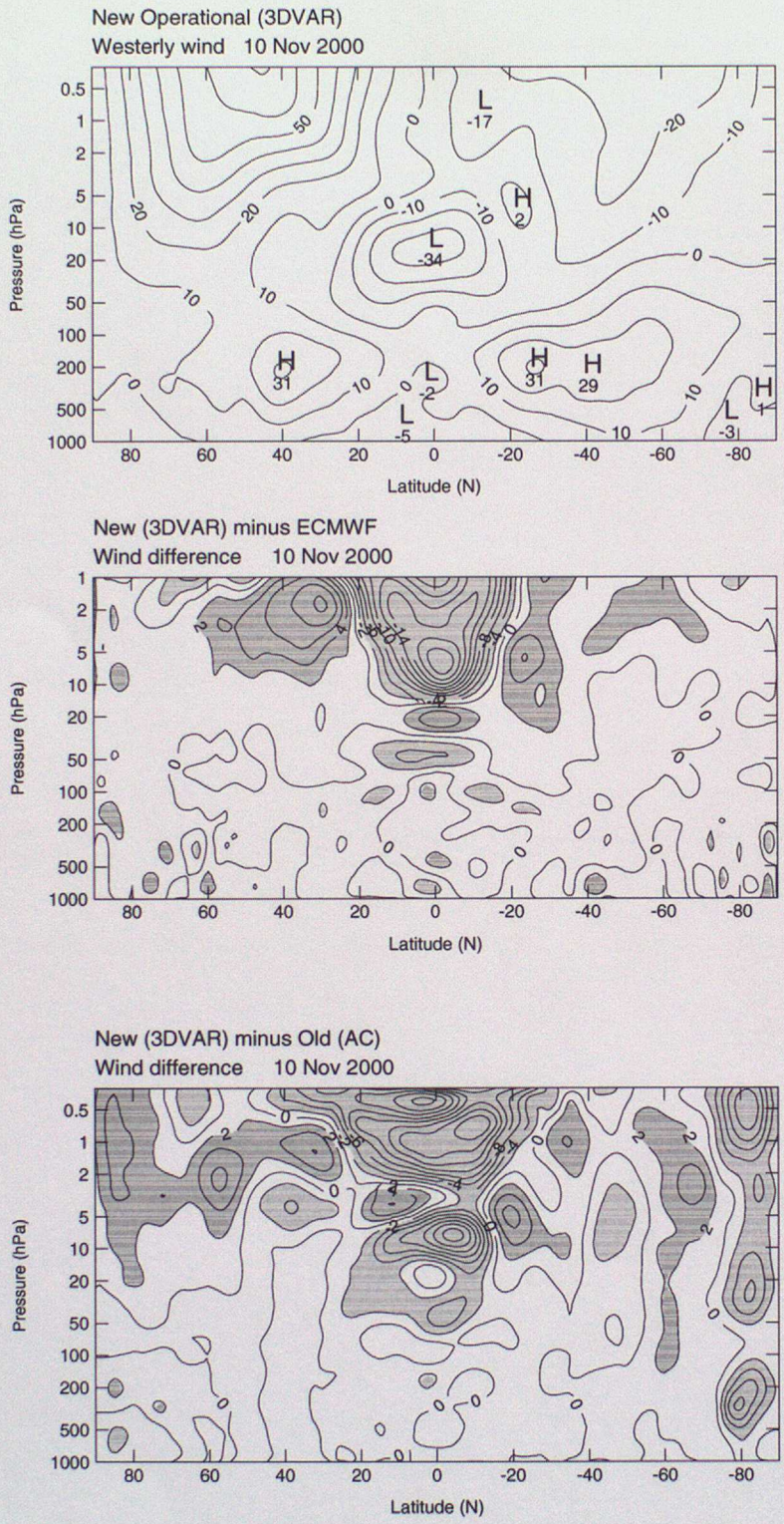


Figure 12. As Fig 11, but for zonal-mean zonal winds